Estimation of co-conditional moments of transmissivity, hydraulic head, and velocity fields

Samuel Hanna & T.-C. Jim Yeh*

Department of Hydrology and Water Resources, The University of Arizona, Tucson, AZ 85721, U.S.A.

An iterative co-conditional Monte Carlo simulation (IMCS) approach is developed. This approach derives co-conditional means and variances of transmissivity (T), head (ϕ), and Darcy's velocity (q), based on sparse measurements of T and ϕ in heterogeneous, confined aquifers under steady-state conditions. It employs the classical co-conditional Monte Carlo simulation technique (MCS) and a successive linear estimator that takes advantage of our prior knowledge of the covariances of T and ϕ and their cross-covariance. In each co-conditional simulation, a linear estimate of T is improved by solving the governing steady-state flow equation, and by updating residual covariance functions iteratively. These residual covariance functions consist of the covariance of T and ϕ and the cross-covariance function between T and ϕ. As a result, the non-linear relationship between T and ϕ is incorporated in the co-conditional realizations of T and ϕ. Once the T and ϕ fields are generated, a corresponding velocity field is also calculated. The average of the co-conditioned realizations of T, ϕ, and q yields the co-conditional mean fields. In turn, the co-conditional variances of T, ϕ, and q, which measure the reduction in uncertainty due to measurements of T and ϕ, are derived. Results of our numerical experiments show that the co-conditional means from IMCS for T and ϕ fields have smaller mean square errors (MSE) than those from a non-iterative Monte Carlo simulation (NIMCS). Finally, the co-conditional mean fields from IMCS are compared with the co-conditional effective fields from a direct approach developed by Yeh et al. [*Water Resources Research*, **32**(1), 85–92, 1996]. © 1998 Elsevier Science Limited.

Keywords: inverse problem, iterative approach, geostatistics, Monte Carlo simulation, heterogeneous aquifers, conditional means.

1 INTRODUCTION

During the past few decades, numerous mathematical models have been developed to solve the inverse problem associated with groundwater systems, given scattered measurements of hydraulic head, ϕ, and transmissivity, T (refer to Yeh[32] and Hanna[14]). One popular method is the minimum-output-error-based approach (e.g. Yeh and Tauxe[31]; Gavalas *et al.*[10]; Neuman and Yakowitz[23]; Neuman[24]; Clifton and Neuman[4]; Cooley[5]; Carrera and Neuman[1,2]; Willis and Yeh[27]). A drawback of this approach is that the solution is non-unique, and the identity of the estimate is often undefined. In other words, using different initial guesses this approach can lead to different results.

Subsequently, it is unclear whether the estimate is a conditional mean, an effective mean, a conditional realization or simply an estimate without any statistical meaning, and the uncertainty associated with the estimated field cannot be addressed properly.

Uniquely identifying the spatial distribution of transmissivity in a heterogeneous aquifer under steady-state flow conditions is an impossible task, unless all the hydraulic heads are known and boundary fluxes are specified. For cases with scattered T and ϕ measurements (or stochastic inverse problems referred to by Yeh *et al.*[30]), a logical inverse approach should adopt the conditional stochastic concept. That is, one should attempt to obtain T and ϕ fields that preserve their observed values at all sample locations and their underlying statistical properties (i.e. the mean and covariance). Furthermore, the estimated T and ϕ

*Corresponding author. E-mail: ybiem@mac.hwr.arizona.edu

fields should satisfy the governing flow equation. According to the conditional probability concept, these fields are conditional realizations of the ensemble. Implicitly, this concept constrains the number of possible T and ϕ fields that can be derived from an inverse model, though many possible realizations of such conditional fields remain. In contrast, the co-conditional mean (the expected value of all possible conditional realizations) is a uniquely defined field. By focusing on estimating the co-conditional mean field, an inverse model, thus, avoids the non-uniqueness problem conceptually.

The geostatistical inverse approach (Kitanidis and Vomvoris[20]; Hoeksema and Kitanidis[16]; Dagan[7]; Rubin and Dagan[25]; Gutjahr and Wilson[12], Hoeksema and Clapp[18]; Kitanidis[21]) is one possible method for estimating co-conditional mean fields of T and ϕ. It relies on the cokriging technique, which takes advantage of the spatial continuity of the natural log of transmissivity field (ln T) and uses the linearized relationship between ln T and ϕ. In cokriging, the unknown f (mean removed ln T) value at a point of interest is estimated by a weighted linear combination of the observed f and h (mean removed ϕ). The weights are determined by conditioning the estimator to be unbiased and to have a minimum variance. Dagan[6,7] and Rubin and Dagan[25] show that when the random f and h fields are jointly Gaussian with known means and covariances, cokriging estimates and covariances are equivalent to the conditional means and covariances for given measurements of f and h.

In general, the cross-covariance function between f and h and the covariance function of h, required in cokriging, are derived from a first-order linearized version of the governing flow equation (Mizell *et al.*[22]; Kitanidis and Vomvoris[20]; and Hoeksema and Kitanidis[16,17]), while the relation between T and ϕ is non-linear, even if the transformation of T (ln T) is adopted. The linearized relation, based on small perturbation theory, will only give reasonable results if the joint distribution of f and h is normal, which is true when the unconditional variance of f (σ_f^2) is small. In cases of large variances of f, the joint distribution of f and h is unlikely to be normal and the use of classical geostatistical techniques is not justified. As in the classical inverse models, inversion of the matrix in the geostatistical inverse approach can suffer from numerical instability problems. The instability is not the problem of the geostatistical approaches, but it arises from the poorly constructed covariance and cross-covariance matrices used in them. Dietrich and Newsam[9] showed that, as the amount of available data increases and the discretization of the system is refined, an ill-conditioned system of equations may arise.

To overcome the above-mentioned problems, Yeh *et al.*[29] proposed an iterative cokriging-like method that combines cokriging and the numerical flow model iteratively. Yeh *et al.*[30] developed a successive linear estimator with the use of a numerical flow model to incorporate the non-linear relationship between T and ϕ. Although they attempted to derive the co-conditional mean fields of T

and ϕ, their approach suffers from theoretical difficulties. Consider the conditional mean flow equation

$$\nabla \left[\langle T_c(x) \rangle \cdot \nabla \langle \phi_c(x) \rangle \right] + \langle \left[t_c(x) \cdot \nabla h_c(x) \right] \rangle = 0 \qquad (1)$$

where subscript 'c' denotes conditioned and $\langle \rangle$ represents the expected value. $\langle T_c \rangle$ and $\langle \phi_c \rangle$ are the co-conditional mean T and ϕ, respectively, while t_c and h_c represent the conditional perturbations of T and ϕ, respectively. Notice that $\langle T_c \rangle$ and $\langle \phi_c \rangle$, themselves alone, in eqn (1) do not satisfy the mass-conservation principle unless the second term in eqn (1) is included. Difficulties in evaluating the second term compel Yeh *et al.*[30] to approximate the conditional mean flow equation by using only the first term in eqn (1). Thus, estimates of the co-conditional mean T and ϕ fields by their approach are merely the co-conditional effective T and ϕ fields that satisfy the continuity equation, and preserve the measured T and ϕ values at sample locations. They claimed that the co-conditional effective T and ϕ fields would be close to the co-conditional mean fields if the heterogeneity is mild. Since their approach attempts to derive the co-conditional mean fields directly, it will be called the direct approach in this paper.

The co-conditional Monte Carlo simulation is another possible approach, which can be utilized to derive the co-conditional means and variances of T and ϕ fields. This simulation approach relies on cokriging and a superposition technique (Journel and Huijbregts[19]) to generate realizations of the conditional T field based on measurements of T and ϕ. The T fields are then used in a numerical model to derive the corresponding realizations of head. By averaging all the realizations, conditional mean fields are thus obtained. Gutjahr *et al.*[13] used a co-conditional method to address the uncertainties in groundwater travel time and paths. Harter and Yeh[15] extended this method to study the effect of conditioning on the uncertainty in predicting flow and solute transport in the vadose zone. Since cokriging is a linear predictor, it does not fully consider the non-linear relationship between T and ϕ fields, and the simulated ϕ fields, based on the co-conditioned T fields, will not necessarily honor the measured head values at the observation locations. This discrepancy will be exacerbated if aquifer heterogeneity increases. Therefore, the T and ϕ fields cannot strictly be the co-conditional fields and variances are not co-conditional variances.

The goal of this paper is to develop a new co-conditional Monte Carlo simulation technique. It determines the co-conditional mean fields of T, ϕ, and q and their conditional variances, without resorting to the conditional mean flow and Darcy's equations. Hence, the problems associated with the direct approach by Yeh *et al.*[30] are avoided. This new co-conditional simulation technique is also different from the previously mentioned Monte Carlo approach: it incorporates the non-linear relationship between T and ϕ fields and can produce head fields that honor the measured values at observation locations. The ability of our new approach is demonstrated through some numerical examples. The proposed technique will be tested

against a hypothetical aquifer. This technique will show the following conclusions: (a) the more observation points we have, the best estimation we can get; (b) the fewer observation errors, the more accurate the estimation; and (c) the less variance of log K, the best estimate we can get. Although the present approach can be applied to any problem, it was limited in this current paper to the study of error-free observations in a hypothetical aquifer. Under the above limitations it was possible to obtain optimum stability conditions, which are not available in the geostatistical approaches. More work is needed to apply the present approach to a real-world case study taking the measurement errors into consideration.

2 PROBLEM FORMATIONS

Consider $\ln T(x)$ of an aquifer to be a stationary stochastic process with a constant unconditional mean, $E[\ln T] = F$, and f is the unconditional $\ln T$ perturbation. The corresponding hydraulic head is given by $\phi(x) = H(x) + h(x)$, where $H = E[\phi]$ and h is the unconditional head perturbation. Suppose we have n_f observed transmissivity values, $f_i^* = (\ln T_i^* - F)$, and n_h observed head values, $h_j^* = (\phi^* - H)$, where $i = 1,..., n_f$ and $j = n_f + 1,..., n_f + n_h$. According to the stochastic concept, many possible realizations of stochastic ϕ and $\ln T$ fields that preserve the observed head and transmissivity values at sample locations exist, and satisfy their underlying statistical properties (i.e. mean and covariance) as well as the governing flow equation. One way to derive these realizations is to use a co-conditional MCS approach similar to that by Gutjahr et al.[13] This co-conditional MCS approach consists of the following steps. First, classical cokriging is performed, based on the observed f_i^* and h_j^*, to construct an approximate conditional f field (mean removed $\ln T$). That is

$$f_{ck}(x_0) = \sum_{i=1}^{n_f} \lambda_{i0}f_i^*(x_i) + \sum_{j=n_f+1}^{n_f+n_h} \mu_{j0}h_j^*(x_j) \qquad (2)$$

where $f_{ck}(x_0)$ is the cokriged f value at location x_0 (the subscript, ck, stands for classical cokriging). The terms λ_{i0} and μ_{j0} are the classical cokriging weights, which can be evaluated as follows:

$$\sum_{i=1}^{n_f} \lambda_{i0}R_{ff}(x_i,x_l) + \sum_{j=n_f+1}^{n_f+n_h} \mu_{j0}R_{hf}(x_j,x_l) = R_{ff}(x_0,x_l)$$

$$l = 1, 2, ..., n_f$$

$$\sum_{i=1}^{n_f} \lambda_{i0}R_{fh}(x_i,x_l) + \sum_{j=n_f+1}^{n_f+n_h} \mu_{j0}R_{hh}(x_j,x_l) = R_{fh}(x_0,x_l)$$

$$l = n_f + 1, n_f + 2, ..., n_f + n_h$$

(3)

where R_{ff} is the unconditional covariance of f (assumed known); R_{hh} and R_{fh} are the unconditional covariance of h and the cross-covariance of f and h, respectively. R_{hh}

and R_{fh} are derived from a first-order numerical approximation (Dettinger and Wilson[8] and Sun and Yeh[26]) for its flexibility for the case of bounded domains and non-stationary problems. Note that F is assumed to be known and $H(x)$ is derived by solving the governing flow equation with the transmissivity equal to $\exp[F]$. Similarly, one can use classical cokriging based on the observed f_i^* and h_j^* to construct a cokriged head field. Next, an unconditional realization of the perturbation of the $\ln T(x)$ field, $f_u(x)$, which maintains the prescribed covariance function is generated. Then, the corresponding unconditional head field, $\phi_u(x) = H(x) + h_u(x)$, is calculated by solving the governing flow equation, with $T_u(x) = \exp(F + f_u(x))$. From these $f_u(x)$ and $h_u(x)$ fields, samples are taken at the same observation locations. Cokriging is again applied using the values of these samples to derive the cokriging estimates, $f_{uk}(x)$ and $h_{uk}(x)$. Finally, conditional realizations of the transmissivity and head fields, $f_c(x)$ and $h_c(x)$, are obtained as follows:

$$f_c(x) = f_{ck}(x) + [f_u(x) - f_{uk}(x)]$$

$$h_c(x) = h_{ck}(x) + [h_u(x) - h_{uk}(x)] \qquad (4)$$

The above-mentioned steps are then repeated, using different seed numbers, to generate different realizations of the conditional fields of $T_c(x)$ and $\phi_c(x)$. Note that $T_c(x) = \exp(F + f_c(x))$ and $\phi_c(x) = H(x) + h_c(x)$ are consistent with the measured values, and $T_c(x)$ retains the same covariance function as that of the true transmissivity field. However, these fields do not necessarily conserve mass since the cokriged fields are linear estimates. This method will be called the non-iterative co- conditional Monte Carlo simulation (NIMCS or Approach I) throughout the rest of the paper.

To circumvent the mass-conservation problem associated with Approach I, our new approach adopts an iterative procedure to improve the co-conditional $f_c(x)$ estimates derived from Approach I. Specifically, once $T_c(x)$ is obtained, it is then used in the governing flow given by eqn (5) along with specified boundary conditions to derive a new head field, ϕ:

$$\nabla[T_c(x)\cdot\nabla\phi(x)] + Q(x) = 0 \qquad (5)$$

In eqn (5), x is the location vector and Q is a deterministic source/sink term. Again, this new head field is an approximate field conditioned on the f and h measurements. It does not necessarily agree with the observed head values at the sample locations due to the assumption of the linear relationship between f and h. At this step, the resultant f and h fields are identical to those by Gutjahr et al.[13] To improve the linearity assumption, a successive linear estimator, similar to the one proposed by Yeh et al.,[30] is adopted to modify the estimate of $T_c(x)$:

$$Y_c^{(r+1)}(x_0) = Y_c^{(r)}(x_0) + \sum_{j=n_f+1}^{n_f+n_h} \omega_{j0}^{(r)}\left[\phi_j^*(x_j) - \phi_j^{(r)}(x_j)\right]$$

(6)

where the ω_{j0} values are the weighting coefficients associated with the estimate at location x_0 with respect to the head measurement at location x_j; r is the iteration index; and Y_c (equal to $F + f_{ck}$, at $r = 0$) is an estimate of the true $\ln T$. The difference between $\ln T$ and Y_c is then denoted by y; $\phi_j^{(r)}$ is the head at the jth location obtained from eqn (5) at iteration r; and ϕ_j^* is the observed head at location j (i.e. $\phi_j^* = Hj + h_j^*$). To determine optimal weighting coefficients that ensure the minimal variance of our new estimate, the minimal mean square error (MSE) criterion is used. That is, $E[(\ln T - Y_c)^2]$ is differentiated with respect to ω_{j0} and the resultant is set to zero. Thus, a system of equations is formed

$$\sum_{j=n_f+1}^{n_f+n_h} \omega_{j0}^{(r)} \epsilon_{hh}^{(r)}(x_j, x_l) = \epsilon_{yh}^{(r)}(x_0, x_l) \quad l = n_f + 1, ..., n_f + n_h$$

(7)

in which ϵ_{hh} represents the covariance of the residual $h = (\phi - \phi^{(r)})$ and ϵ_{yh} represents the cross-covariance between residuals h and $y = (\ln T - Y_c^{(r)})$. If ϵ_{yh} and ϵ_{hh} are given, the ω values can be determined by solving eqn (7). With new values, eqn (6) is employed to update our estimate of Y_c. The covariance and cross-covariance, ϵ_{hh} and ϵ_{yh}, required in eqn (7) are approximated by a first-order analysis for a finite element groundwater-flow model as described by Equation (12) in Yeh et al.[30] The covariance of the residual y (ϵ_{yy}) is determined at each iteration as follows:

$$\epsilon_{yy}^{(1)}(x_0, x_k) = R_{ff}(x_0, x_k) - \sum_{i=1}^{n_f} \lambda_{i0} R_{ff}(x_0, x_i)$$

$$- \sum_{j=n_f+1}^{n_f+n_h} \mu_{j0} R_{fh}(x_k, x_j) \text{ for } r = 0$$

(8)

$$\epsilon_{yy}^{(r+1)}(x_0, x_k) = \epsilon_{yy}^{(r)}(x_0, x_k) - \sum_{i=n_f+1}^{n_f+n_h} \omega_{i0}^{(r)} \epsilon_{yh}^{(r)}(x_k, x_i)$$

for $r \geq 1$

where $k = 1, 2, ..., N$, λ and μ are weighting coefficients, and R_{ff} is the given unconditional covariance function of f. After updating $Y_c^{(r)}(x)$, the flow given by eqn (5) is solved again with the updated $Y_c^{(r)}(x)$ for a new head field, $\phi(r)$. This iterative procedure continues until the absolute difference in σ_f^2 (the variance of the estimated $\ln T$ field) between two successive iterations is smaller than a prescribed tolerance. If the criterion is not met, new ϵ_{yh} and ϵ_{hh} are evaluated again, and eqn (7) is solved to obtain a new set of weights. These weights are then used in eqn (6) with new values of $(\phi_j^* - \phi_j^{(r)})$ to obtain a new estimate, $Y_c(x)$. We refer to this new approach as an iterative co-conditional Monte Carlo simulation technique (IMCS) and it will be labeled as Approach II in the following numerical examples. Based on Approach II, different starting $T_c(x)$ fields will result in different $Y_c(x)$ and $\phi_c(x)$ fields after the iteration process. Averaging all realizations

of $Y_c(x)$ and $\phi_c(x)$ fields yields the co-conditional mean transmissivity and head fields. In turn, the co-conditional variances of T and ϕ are evaluated.

The concept of our new approach is different from that of the direct approach by Yeh et al.[30] although the algorithm is similar. The direct approach attempts to derive the co-conditional mean fields by using the cokriged T as a starting estimate of the conditional mean T field, then it applies an iterative procedure to update the estimate. In theory, their approach requires them to solve the conditional mean flow equation. However, difficulties in evaluating the second term in eqn (1) forced them to approximate the conditional mean flow equation by using eqn (5). In contrast, our new approach starts with estimates of co-conditional realizations of T fields, and then solves the exact governing flow equation of conditional realizations during the iteration. Thus, the use of the approximate mean equation is avoided and the co-conditional variance can be evaluated directly without the use of the first-order approximation.

3 EXAMPLES

Application of inverse models to any field problems, generally yields inconclusive results unless the field site is fully characterized with detailed data. No such field site exists, and testing of our stochastic approach will rely on a hypothetical heterogeneous aquifer where the hydraulic properties and boundary conditions are known exactly. This aquifer is assumed to be 40 m $\times$ 40 m in dimension and is discretized into 1600 elements (each 1 m $\times$ 1 m). Each element is assigned a $\ln T$ value using a random field generator developed by Gutjahr.[11] This generated $\ln T$ field has zero mean and an exponential correlation structure with variance, σ_f^2, of 3.24, and anisotropic correlation scales ($x = 6$ m and $\lambda_y = 2$ m in the x and y directions, respectively). The upper and lower sides of the aquifer are assigned as no-flow boundaries, and the left and right sides are prescribed head boundaries with head values 10.4 m and 10.0 m, respectively. In addition, a pumping well with a constant discharge ($Q_w = 3$ m^3/m^2/s) is located at a point (18 m, 23 m). With the generated transmissivity field and the prescribed conditions, a finite element model based on the approach of Yeh et al.[28] is then used to derive the hydraulic head distribution at nodal points. The head values at the four nodes of an element are averaged to represent the head at the center of the element. Using these head and transmissivity fields, the corresponding Darcy's velocity field is then calculated based on Darcy's law. Thereafter, these generated transmissivity, head, and velocity fields will be called the true fields (Fig. 1(a)–(c)).

From the true transmissivity and head fields, forty-two transmissivity ($n_f = 42$) and sixty-three head values ($n_h = 63$) are sampled at uniformly distributed locations (circles in Fig. 1(a) and (b)). These transmissivity and head values are considered as our measurements. Note that we choose fewer
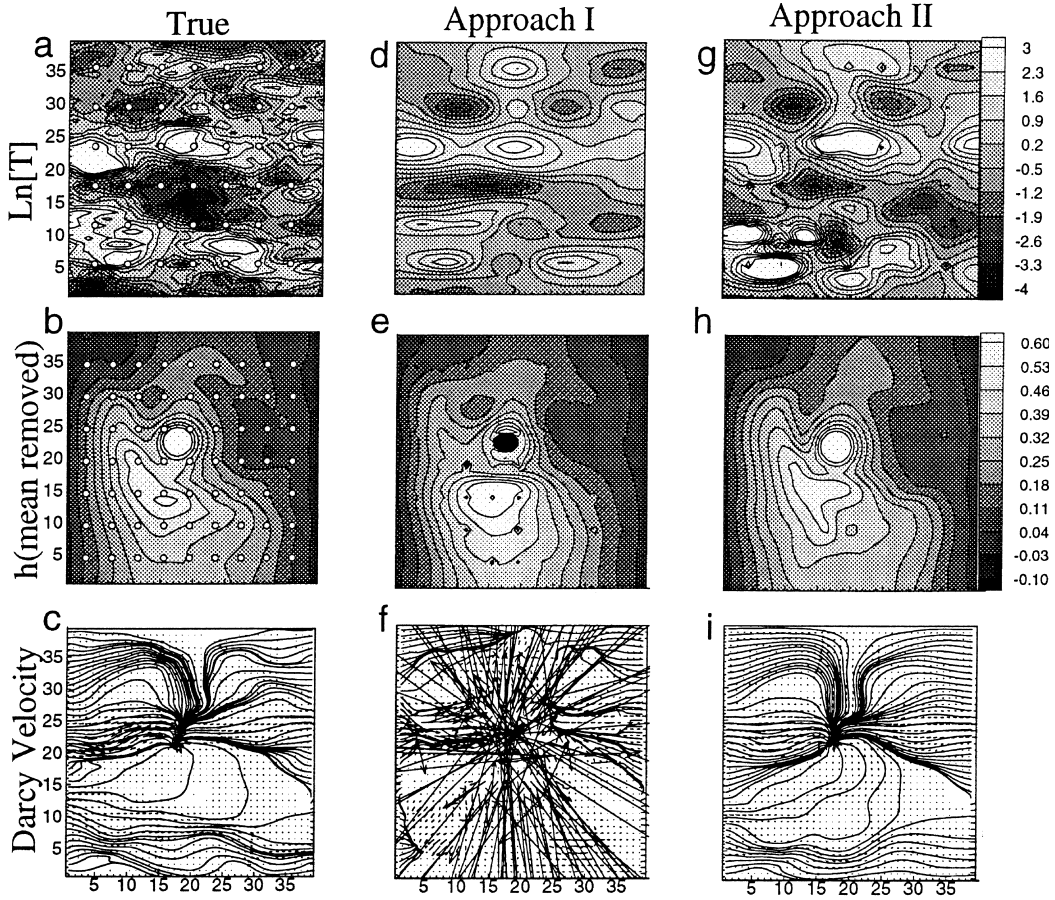
Fig. 1. Comparisons of the true transmissivity, head, and Darcy's velocity and those derived from Approaches I and II.

transmissivity observations than the head ones for practical purposes and to show how our approach will perform. Assuming that the covariance function of f and boundary conditions are known exactly, 400 realizations of f and h fields conditioned on the measurements are produced using Approaches I and II. Darcy's law is employed to derive the corresponding velocity fields. By averaging these realizations, co-conditional mean f, h, and velocity fields are determined, which are compared with the true fields using the following criteria:

$$P_1 = \frac{1}{N} \sum_{i=1}^{N} (Z_{0i} - Z_{ei}) \text{ and } P_2 = \frac{1}{N} \sum_{i=1}^{N} (Z_{0i} - Z_{ei})^2 \quad (9)$$

In eqn (9), Z_{0i} and Z_{ei} are the true and estimated co-conditional mean transmissivity or head values at the ith location, respectively, and N is the total number of elements. P_1 is a measure of the bias, and P_2 is the MSE of the estimates.

4 RESULTS AND DISCUSSION

A visual illustration of the performance of the two approaches is provided in Fig. 1. In general, the f fields derived from Approach I (Fig. 1(d)) and Approach II

(Fig. 1(g)) are smoother than the true f field. This is attributed to the nature of conditional expectation with a limited number of observations. Both approaches depict the general spatial structure of the true f field, while the f field from Approach II reveals more details, since it incorporates the non-linearity between f and h in the iterative process. In addition, each realization of the co-conditional f and h fields, generated by Approach II, satisfies the governing flow equation during each iteration. As a result, the conditional head field from Approach II is improved successively, not only at the observed locations but also at their vicinity. The resultant head field of Approach II is much closer to the true one than that of Approach I (Fig. 1(h) and (e)). Some spurious kinks appear in the head field from Approach I, which can be attributed to the linear operation of cokriging and the superposition technique.

As expected, the conditional mean velocity field from Approach II is smoother than the true one (Fig. 1(i) and (c)). The velocity resulting from Approach I (Fig. 1(f)) suffers from anomalous sources and sinks due to the mass-conservation problem mentioned previously. Again, since each conditional realization of f and h fields, derived from Approach II, satisfies the governing flow equation, these realizations yield mass-conservative velocity fields using Darcy's law. The average of all the
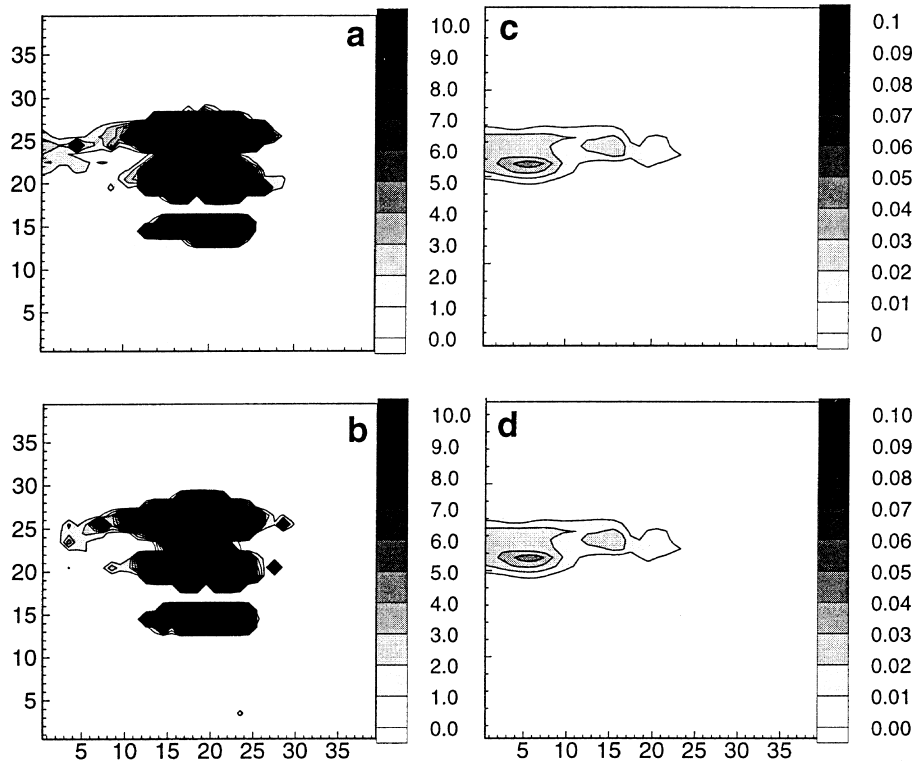
Fig. 2. Contours of the conditional standard deviation of (a) q_x and (b) q_y for Approach I; contours of the conditional standard deviation of (c) q_x and (d) q_y for Approach II.
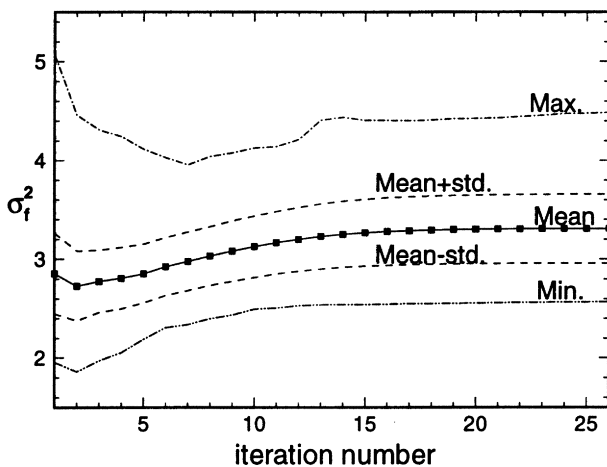
velocity fields leads to the conditional mean velocity which also satisfies the mass-conservation principle.

Contour maps of the standard deviation of the velocity in x and y directions (q_x and q_y, respectively) are shown in Fig. 2(a) and (b) for Approach I and in Fig. 2(c) and (d) for Approach II. According to these figures, Approach II produces conditional velocity fields with smaller conditional standard deviations (less uncertainty). Large values of the standard deviation near the pumping well (Fig. 2(a) and (b)) reflect the effect of strong non-linearity caused by the non-uniform flow regime, which was not considered in


Fig. 3. Mean, maximum and minimum of the variance of the estimated f field as a function of the number of iterations.

Approach I. This is not the case with Approach II, demonstrating that the iterative procedure can handle the non-linearity between f and h.

To demonstrate the convergence of our iterative procedure, the mean, the deviation around the mean, and the range of σ_f^2 of 400 conditional f realizations are plotted as a function of the iteration number in Fig. 3. Fig. 3 shows approximately a sort of stability of the calculated variance of log K as the number of iterations increases. The standard deviation remains almost constant throughout the iteration process, indicating that at the end of the iterative process each conditional realization of f field maintains nearly the same value of σ_f^2 as the initial unconditional f field. In other words, the iterative method seems to converge to the right optimum. The standard deviation around the mean in the Monte Carlo simulation is attributed to the effect of the small domain size used in the study.

Fig. 4(a) shows the change of the variances of the estimated co-conditional mean f field, Σ_f^2, by the two approaches with the number of realizations. It can be seen that Σ_f^2 of Approach II converges to a value that is larger than that of Approach I, and is smaller than that for the true field. This result is consistent with the previous discussion of Fig. 1: the incorporation of the non-linearity between f and h reveals more variability and the co-conditional mean field is smoother than the true field. MSE values of conditional mean f and h fields evaluated at different numbers of realizations are depicted in Fig. 4(b) and (c), respectively. Based on these results, one can conclude that (a) variances
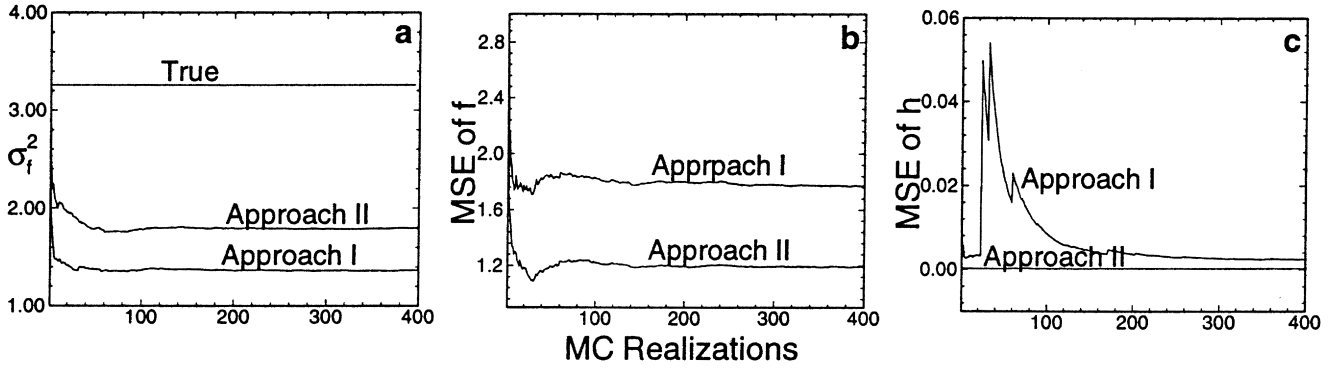
Fig. 4. (a) True transmissivity field variance and those from Approaches I and II. (b) MSE of transmissivity estimates of Approaches I and II. (c) MSE of the head estimates from Approaches I and II as a number of Monte Carlo realizations.

and MSE values reach steady values after about 150 realizations, which is far less than the number of realizations for the unconditional MCS (in agreement with results of Harter and Yeh[15]); (b) the MSE for hydraulic head stabilizes very rapidly and approaches zero after a few realizations for Approach II, while it takes many more realizations to stabilize in Approach I; (c) the MSE values for the conditional means of f and h in Approach II are smaller than those in Approach I.

As mentioned previously, our iterative co-conditional Monte Carlo simulation technique does not require the use of the conditional mean flow equation. The average of the realizations of the co-conditional Monte Carlo simulation should theoretically be better than that derived from the direct approach by Yeh et al.[30] Nevertheless, the conditional effective means of f, h, q_x and q_y fields from the direct approach are in good agreement with those from Approach II (Fig. 5), suggesting that the direct approach closely approximates the conditional mean fields. This result seems plausible because the effect of the term involving products of perturbations in the mean flow equation diminishes as the number of observations of f or h increases. This conclusion is important since the direct approach does not need to conduct a large number of Monte Carlo simulation runs, resulting in a substantial saving in CPU time.
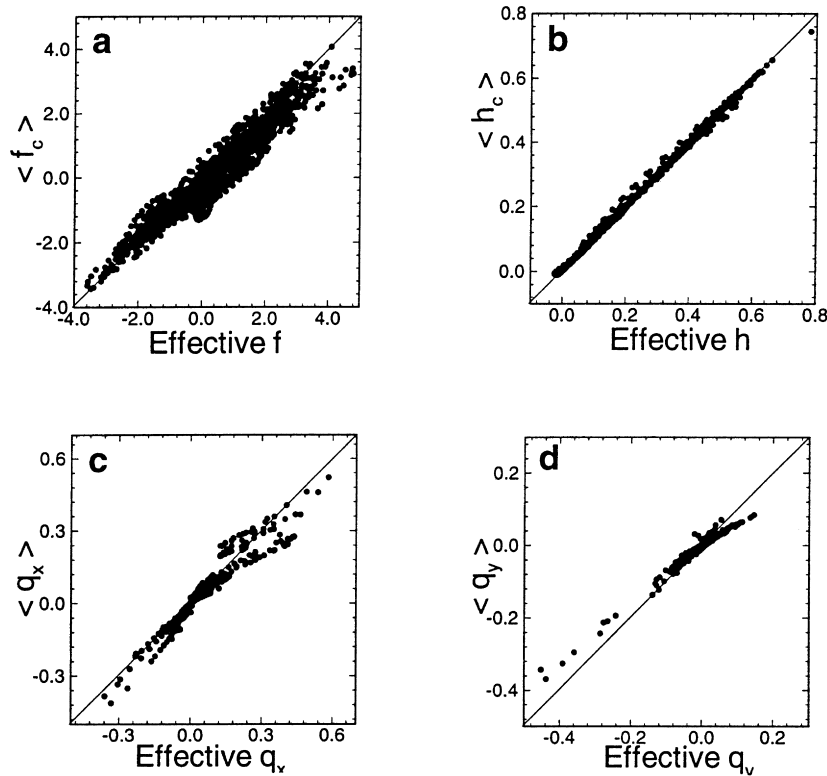
Fig. 5. (a) Co-conditional transmissivity field from Approach II versus the effective transmissivity field by the direct approach. (b) Co-conditional head field from Approach II versus the effective head field from the direct approach. (c) Co-conditional Darcy's velocity field q_x from Approach II versus the effective q_x. (d) Co-conditional Darcy's velocity field q_y from Approach II versus the effective q_y.

Finally, the inverse method based on the minimal-output-error approach (such as Carrera and Glorioso[3]) can be used to replace our successive linear estimator in our iterative co-conditional Monte Carlo simulation algorithm. However, the approximated co-conditional f field derived from Approach I is essential to be used as the starting f field in the minimal-output-error approach. The approximate field is close to the correct co-conditional realization. Thus a global minimum, instead of local minima, can be obtained. Such an approach, in our opinion, should result in the same co-conditional mean and variance fields as those by Approach II. In this way, the identity problem associated with the classical minimal-output-error approach is eliminated.

Our technique and the one developed by Kitanidis[21] are conceptually identical. However, the total number of equations solved using our technique is equal to the number of head measurements plus one. On the other hand, the method of Kitanidis[21] requires solving a system of equations of a size that equals the number of head measurements plus the number of transmissivity measurements.

5 CONCLUSION

Our proposed IMCS approach is an extension of that by Yeh *et al.*[30] It attempts to include the non-linear relationship between f and h through successive linear approximations. A hypothetical aquifer was used to demonstrate the ability of the approach, and we show that iterative co-conditional Monte Carlo simulation is better than non-iterative co-conditional Monte Carlo simulation. The iterative approach can produce realizations of transmissivity and head fields that agree with the observations at measurement locations even in highly heterogeneous aquifers under non-uniform flow conditions. In addition, iterative co-conditional Monte Carlo simulation can yield mass-conservative co-conditional mean velocities and mass-conservative conditional variances, which are crucial in the stochastic analysis of solute transport in heterogeneous aquifers. Since our approach is a Monte Carlo simulation approach, it requires significant computational effort. We believe this shortcoming can be overcome in the future by the rapid advances in computing technology. It is not our intention to demonstrate the ability of our technique for solving real-world problems although our approach is not limited to zero observation error. For field problems, where measurement errors and other unknown factors may play important roles, more theoretical development is still needed and the ability of our iterative approach remains to be tested.

ACKNOWLEDGEMENTS

REFERENCES

1. Carrera, J. & Neuman, S. P. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. *Water Resour. Res.*, 1986, **22**(2), 199–210.
2. Carrera, J. & Neuman, S. P. Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability, and solution algorithms. *Water Resour. Res.*, 1986, **22**(2), 211–227.
3. Carrera, J. & Glorioso, L. On geostatistical formulation of the groundwater flow inverse problem. *Adv. Water Resources*, 1991, **14**(5), 273–283.
4. Clifton, P. M. & Neuman, S. P. Effects of kriging and inverse modeling on conditional simulation of the Avra valley aquifer in southern Arizona. *Water Resour. Res.*, 1982, **18**(4), 1215–1234.
5. Cooley, R. L. Incorporation of prior information of parameters into nonlinear regression groundwater models, 1. Theory. *Water Resour. Res.*, 1982, **18**(4), 965–976.
6. Dagan, G. Stochastic modeling of groundwater flow by unconditional and conditional probabilities, 1. Conditional simulation and the direct problem. *Water Resour. Res.*, 1982, **18**(4), 813–833.
7. Dagan, G. Stochastic modeling of groundwater flow by unconditional and conditional probabilities: the inverse problem. *Water Resour. Res.*, 1985, **21**(1), 65–72.
8. Dettinger, M. D. & Wilson, J. L. First order analysis of uncertainty in numerical models of ground water flow, Part 1. Mathematical development. *Water Resour. Res.*, 1981, **17**(1), 149–161.
9. Dietrich, C. R. & Newsam, G. N. A stability analysis of the geostatistical approach to aquifer transmissivity identification. *J. of Stochastic Hydraul.*, 1989, **3**, 293–316.
10. Gavalas, G. R., Shah, P. C. & Seinfeld, J. H. Reservoir history matching by Bayesian estimation. *Trans. AIME*, 1976, **261**, 337–350.
11. Gutjahr, A. L., Fast Fourier transforms for random field generation, Project Report for Los Alamos Grant, Department of Mathematics, New Mexico Institute of Mining and Technology, April 1989.
12. Gutjahr, A. L. & Wilson, J. L. Co-kriging for stochastic models. *Transport in Porous Media*, 1989, **4**(6), 585–598.
13. Gutjahr, A. L., Bullard, B., Hatch, S. & Hughson, L. Joint conditional simulations and the spectral approach for flow modeling. *J. of Stoch. Hydrol.*, 1994, **8**(1), 79–108.
14. Hanna, S., An iterative Monte Carlo technique for estimating conditional means and variances of transmissivity and hydraulic head fields. Ph.D. dissertation, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ, 1995.
15. Harter, T. & Yeh, T.-C. J. Conditional stochastic analysis of solute transport in heterogeneous, variably saturated soils. *Water Resour. Res.*, 1996, **32**(6), 1597–1610. (in press).
16. Hoeksema, R. J. & Kitanidis, P. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resour. Res.*, 1984, **20**(7), 1003–1020.
17. Hoeksema, R. J. & Kitanidis, P. K. Prediction of transmissivities, heads, and seepage velocities using

mathematical modeling and geostatistics. *Adv. Water Resour.*, 1989, **12,** 90–101.

18. Hoeksema, R. J. & Clapp, R. P., Calibration of groundwater flow models using Monte Carlo simulations and geostatistics. In *ModelCARE 90: Calibration and Reliability in Groundwater Modeling*. IAHS Publication No. 195, 1990, pp. 33–42.

19. Journel, A.G. & Huijbregts, Ch. J., *Mining Geostatistics*. Academic Press, London, 1978, p. 600.

20. Kitanidis, P. & Vomvoris, E. A geostatistical approach to the inverse problem in groundwater modeling (steady-state) and one-dimensional simulation. *Water Resour. Res.*, 1983, **19,** 677–690.

21. Kitanidis, P. K. Quasilinear geostatistical theory for inversing. *Water Resour. Res.*, 1995, **31**(10), 2411–2419.

22. Mizell, S. A., Gutjahr, A. L. & Gelhar, L. W. Stochastic analysis of spatial variability in two-dimensional steady groundwater flow assuming stationary and nonstationary heads. *Water Resour. Res.*, 1982, **18**(4), 1053–1067.

23. Neuman, S. P. & Yakowitz, S. A statistical approach to the inverse problem of aquifer hydrology, 1. Theory. *Water Resour. Res.*, 1979, **15**(4), 845–860.

24. Neuman, S. P. A statistical approach to the inverse problem of aquifer hydrology, 3. Improved solution method and added perspective. *Water Resour. Res.*, 1980, **16**(2), 331–346.

25. Rubin, Y. & Dagan, G. Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 1. Theory. *Water Resour. Res.*, 1987, **23**(7), 1185–1192.

26. Sun, N.-Z. & Yeh, W. W.-G. A stochastic inverse solution for transient groundwater flow: parameter identification and reliability analysis. *Water Resour. Res.*, 1992, **28**(12), 3269–3280.

27. Willis, R., & Yeh, W.-G., *Groundwater Systems Planning and Management*. Prentice Hall, Englewood Cliffs, NJ, 1987.

28. Yeh, T.-C.J., Srivastava, R., Guzman, A. & Harter, T. A numerical model for water flow and chemical transport in variably saturated porous media. *Groundwater*, 1993, **31,** 634–644.

29. Yeh, T.-C.J., Gutjahr, A. L. & Jin, M. An iterative cokriging-like technique for groundwater flow modeling. *Groundwater Water*, 1995, **33**(1), 33–41.

30. Yeh, T.-C.J., Jin, M. & Hanna, S. An iterative stochastic inverse method: conditional effective transmissivity and hydraulic head fields. *Water Resour. Res.*, 1996, **32**(1), 85–92.

31. Yeh, W.W.-G. & Tauxe, G.W. Optimal identification of aquifer diffusivity using quasilinearization. *Water Resour. Res.*, 1971, **7**(4), 955–962.

32. Yeh, W.W.-G. Review of parameter identification procedure in groundwater hydrology: the inverse problem. *Water Resour. Res.*, 1986, **22**(2), 95–108.